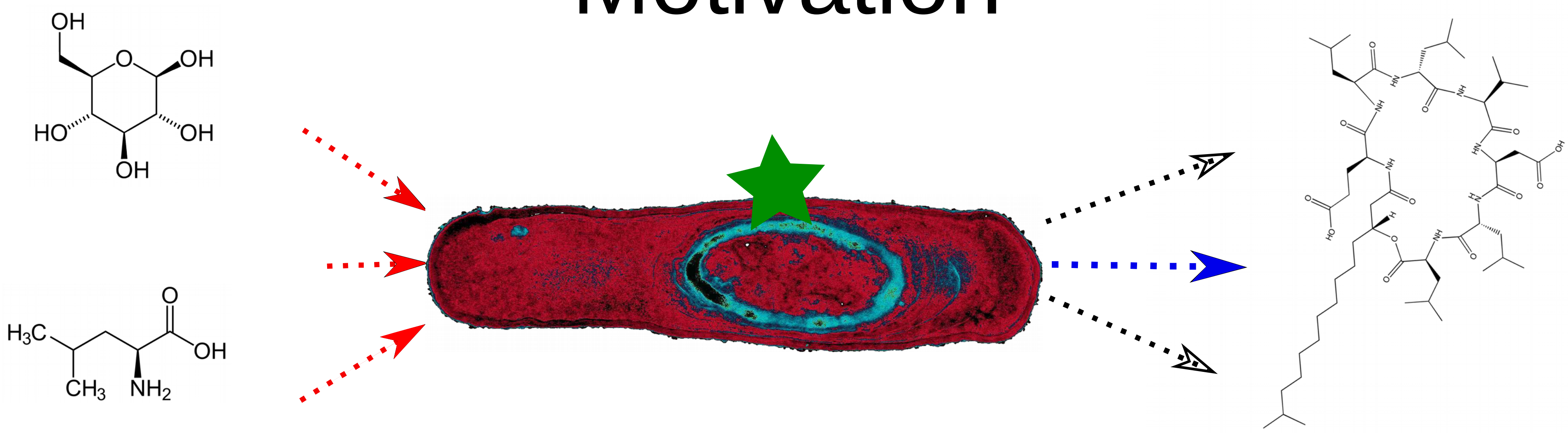
Elementary modes refine abstract interpretation of reaction networks with partial kinetic information

Emilie Allart

Joachim Niehren - Cristian Versari
Cédric Lhoussaine

Motivation



optimal feeding + gene knockout → surfactin overproduction

Want model based prediction of changes that may lead to overproduction

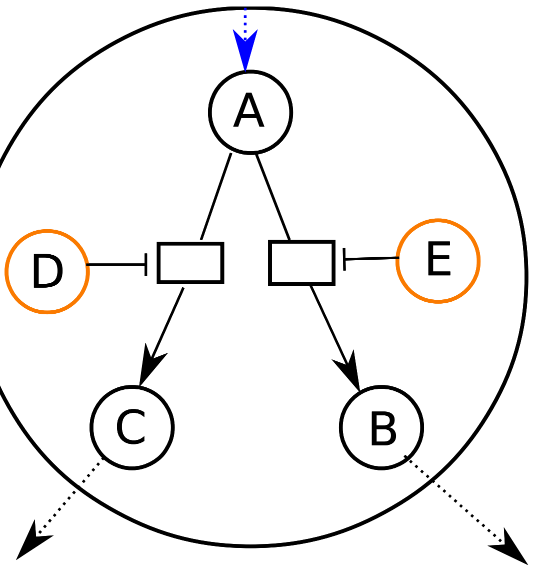
F. Coutte, J. Niehren, D. Dhali, M. John, C. Versari, and P. Jacques, “Modeling Leucine’s Metabolic Pathway and Knockout Prediction Improving the Production of Surfactin, a Biosurfactant from Bacillus Subtilis,” Biotechnology Journal, vol. 10, pp. 1216–34, Aug. 2015.

Context

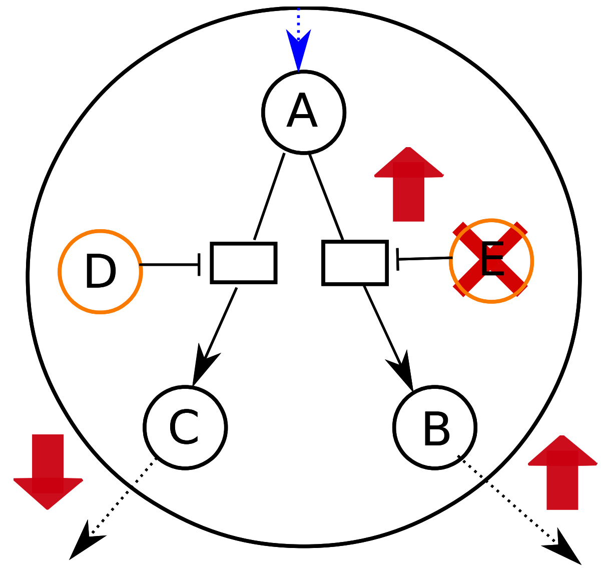
- We have :
  - Structural knowledge on reactions available
  - Partial kinetic information : inhibition & enzymes
  - No temporal information : only steady state
  - No quantitative information

- We use :
  - Flux balance analysis (*Jeffrey D. Orth, Ines Thiele & Bernhard Palsson*)
  - Constraint-based approaches (*Stephen Klamt & Jörg Stelling*)
  - Qualitative reasoning by abstract interpretation (*Joachim Niehren & Cristian Versari*)

Idea of qualitative reasoning



Idea of qualitative reasoning



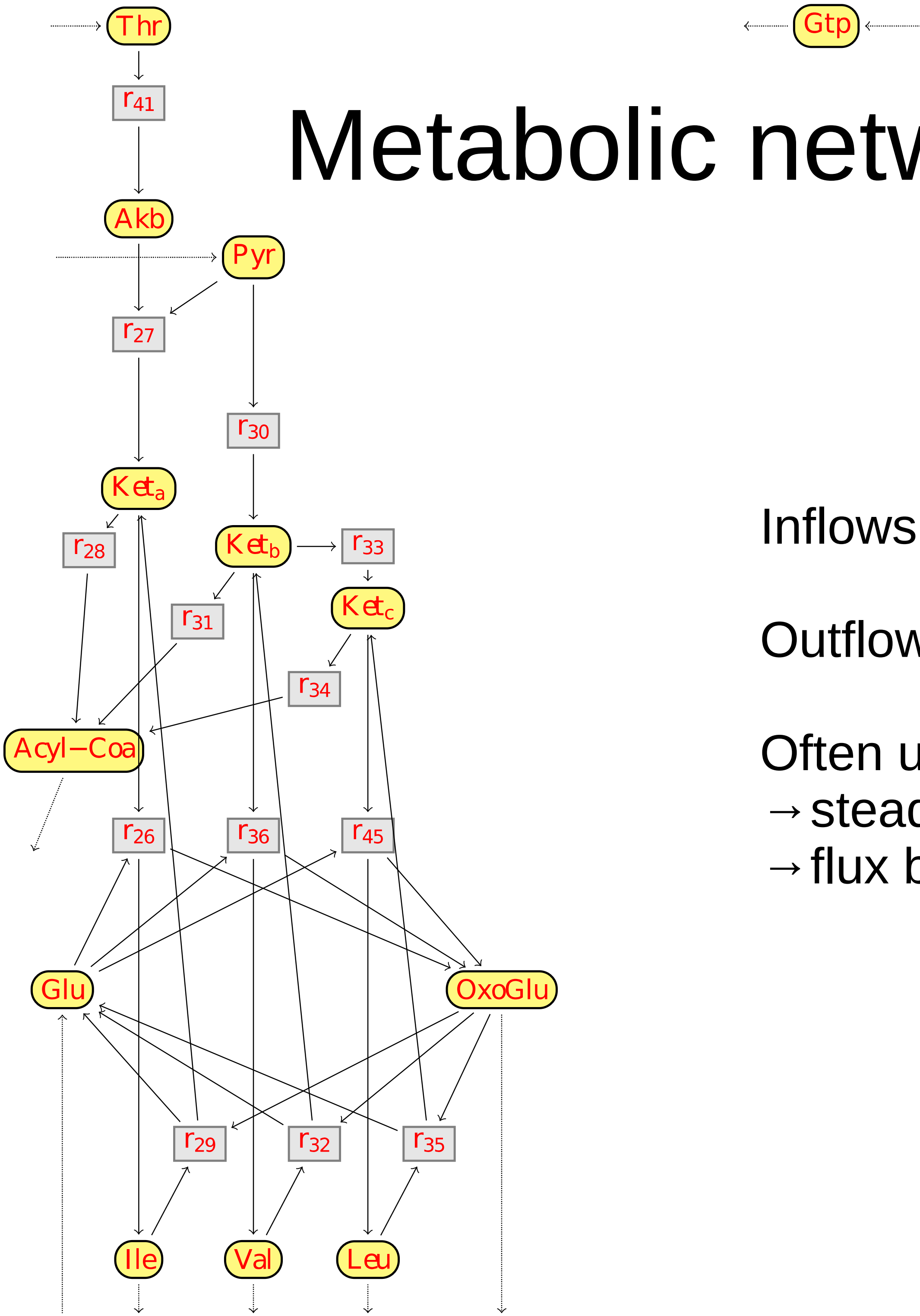
Plan

- Background
  - Modeling language for reaction networks with partial kinetic information
  - Prediction by abstract interpretation
- Over-approximation problem
- Additional constraints from 'elementary mode'

Plan

- Background
  - **Modeling language for reaction networks with partial kinetic information**
  - Prediction by abstract interpretation
- Over-approximation problem
- Additional constraints from 'elementary mode'

Metabolic networks

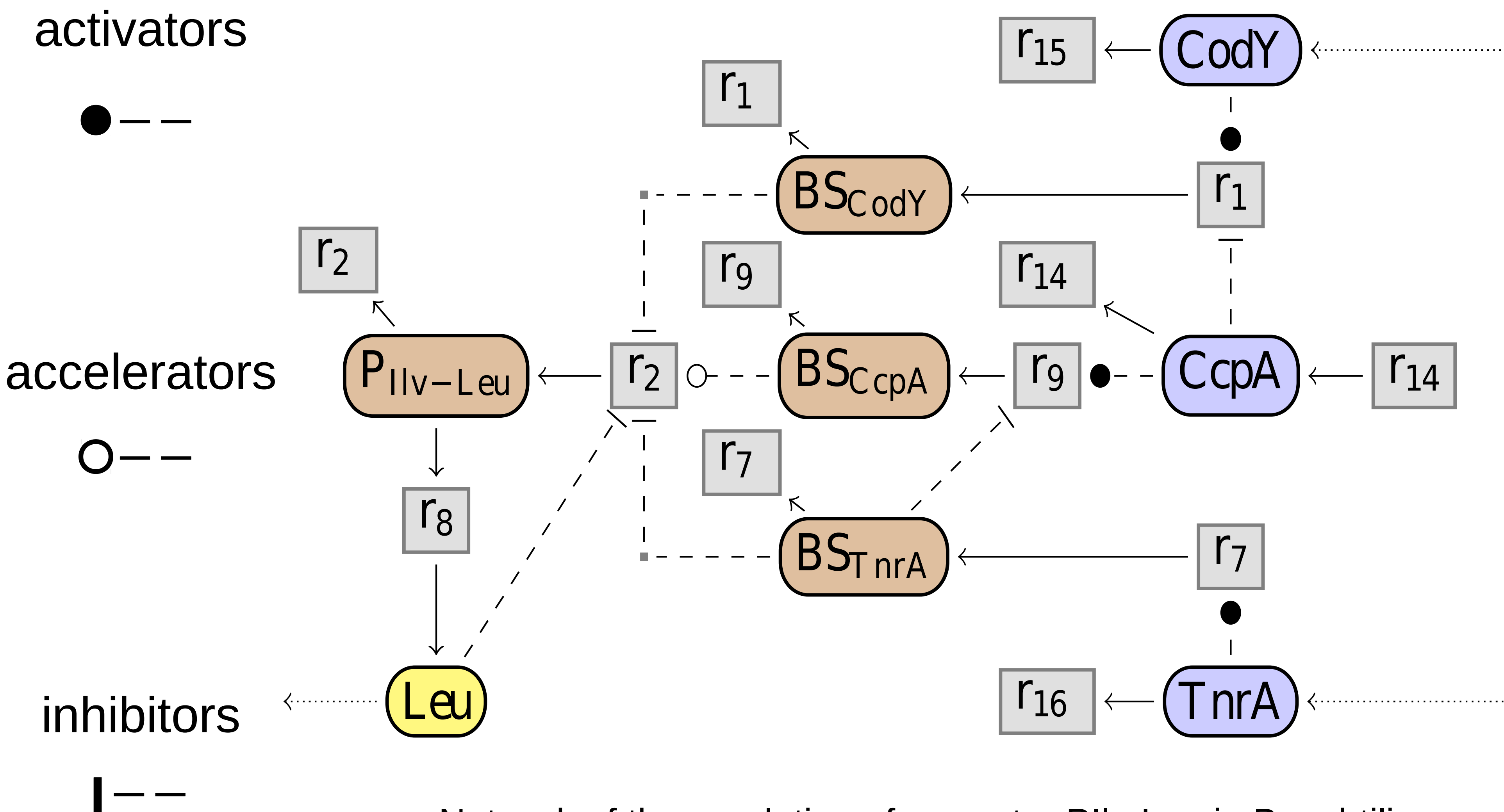


Inflows: externally controlled

Outflows: target

Often unknown kinetic constants
→ steady state semantics
→ flux balance

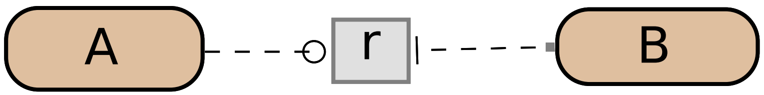
Regulation



Network of the regulation of promoter PIlv-Leu in B. subtilis

Semantics

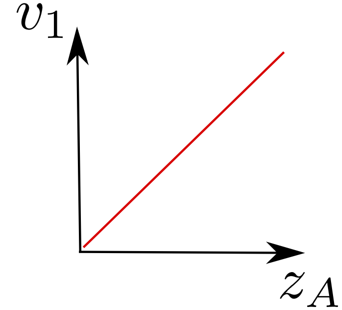
- General kinetic function “exp” :



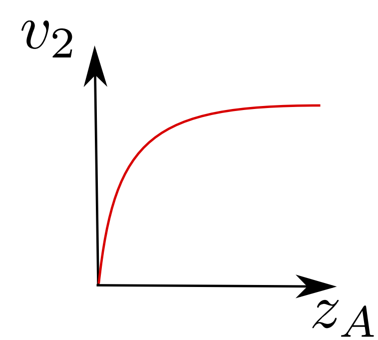
$$\exp\left(inh : z_B, acc : z_A\right) = \frac{z_A}{1 + z_B}$$

- Similarity
  → kinetic functions with same qualitative behaviour
  → all kinetic functions are similar to expressions kinetics
    - monotone in activators and substrates
    - anti-monotone in inhibitors
- Example :

$$ma_k \simeq mm_k$$



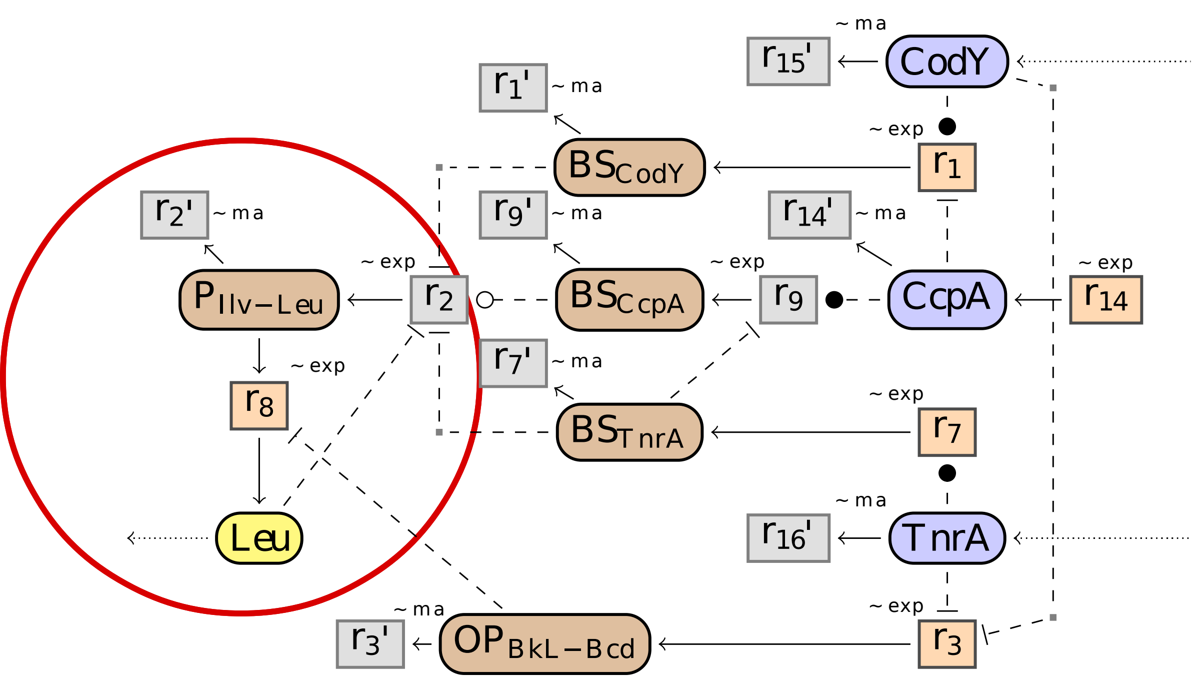
$v_1 = ma_k(subs : a, b)$



$v_2 = mm_k(subs : a)$

Steady State Equations

x_S	Inflow
y_S	Outflow
z_S	Species concentration
v_r	Reaction speed



$$v_{r_2} = \exp^{(2)}(\text{inh}{:}z_{BS_{CodY}},\ \text{acc}{:}z_{BS_{CcpA}},\ \text{inh}{:}z_{Leu},\ \text{inh}{:}z_{BS_{TnrA}})$$

$$v_{r_8} = \exp^{(8)}(\text{subs}{:}z_{P_{Ilv-Leu}})$$

$$v_{r_{2'}} = \text{ma}^{(2)}(\text{subs}{:}z_{P_{Ilv-Leu}})$$

$$y_{Leu} = \text{ma}^{(0)}(\text{subs}{:}z_{Leu})$$

(Leu) $\quad v_{r_8} = y_{Leu}$

$(P_{Ilv-Leu})$ $\quad v_{r_2} = v_{r_2}' + v_{r_8}$

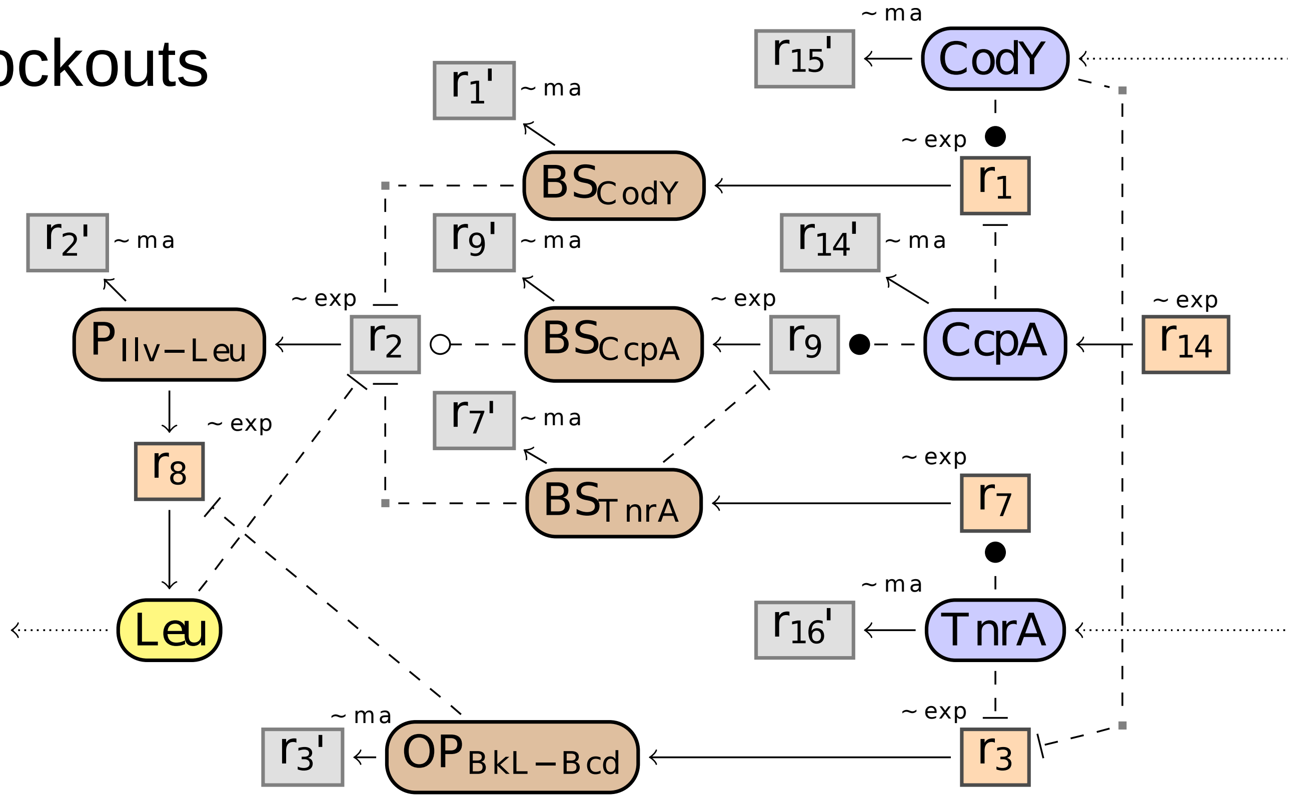
Plan

- Background
  - Modeling language for reaction networks with partial kinetic information
  - **Prediction by abstract interpretation**
- Over-approximation problem
- Additional constraints from 'elementary mode'

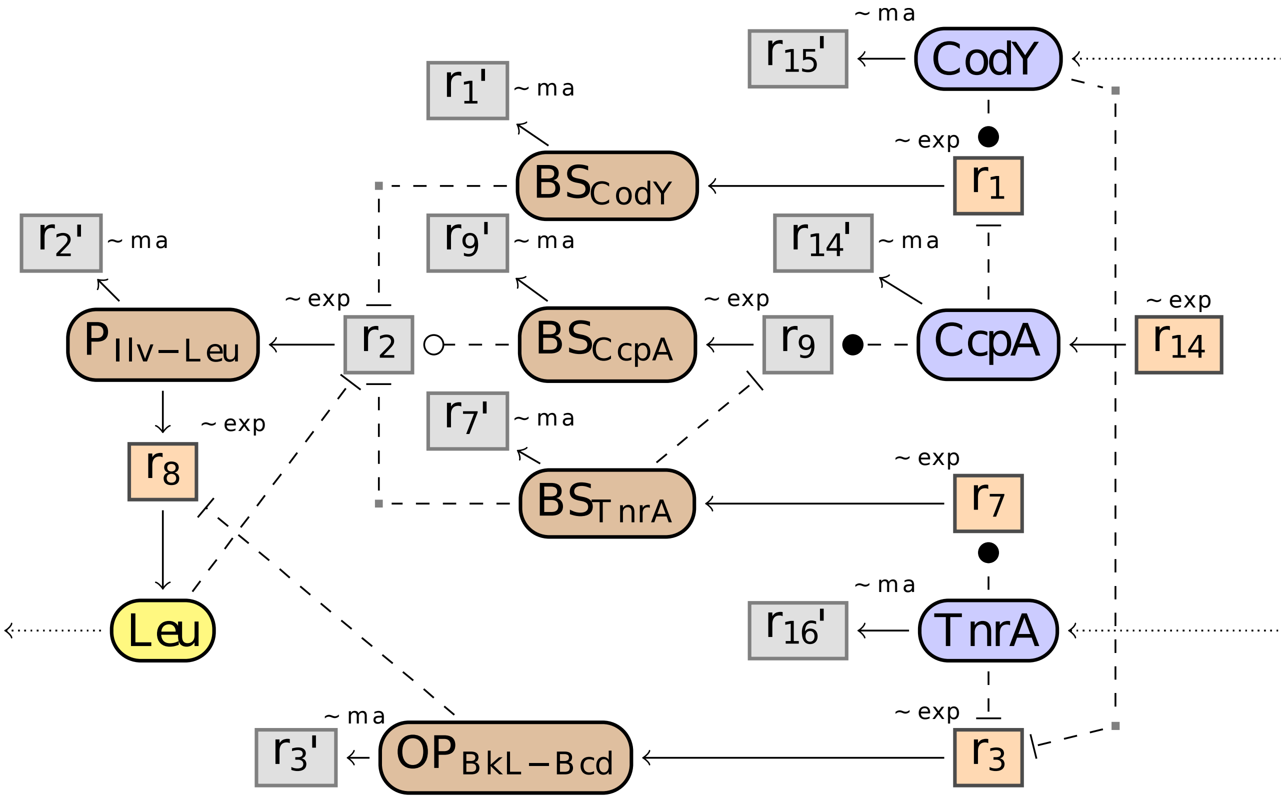
Network changes

Allowed changes:

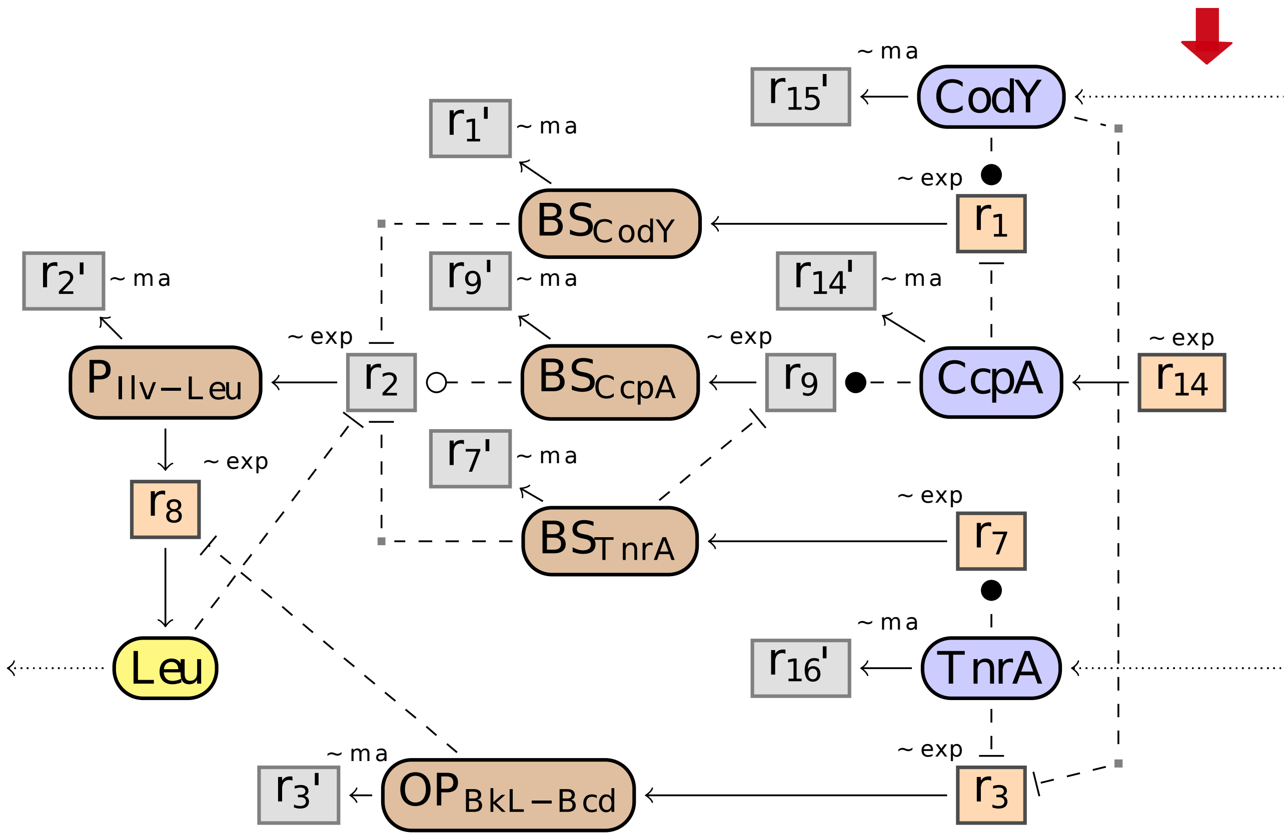
- inflow increase/decrease
- reaction knockouts



Qualitative reasoning

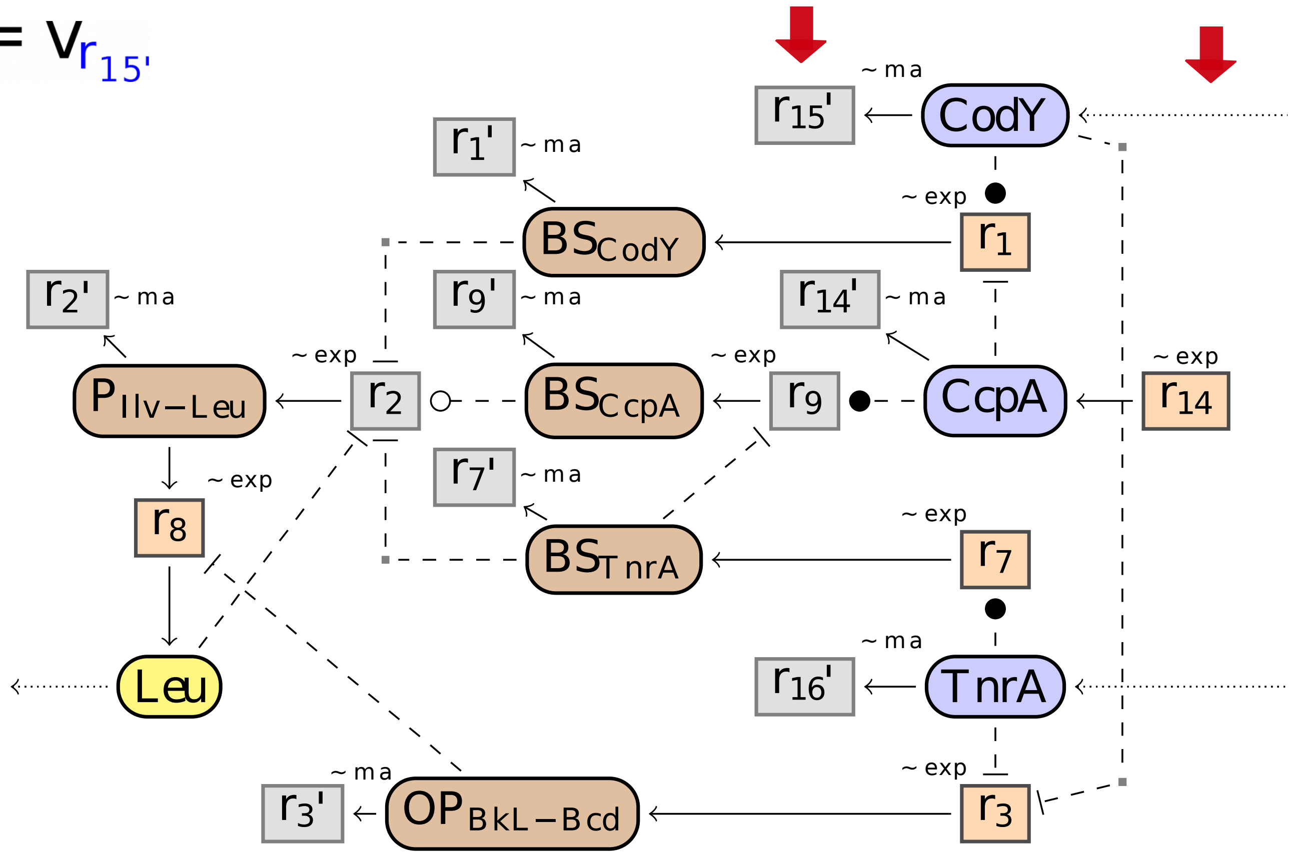


Qualitative reasoning



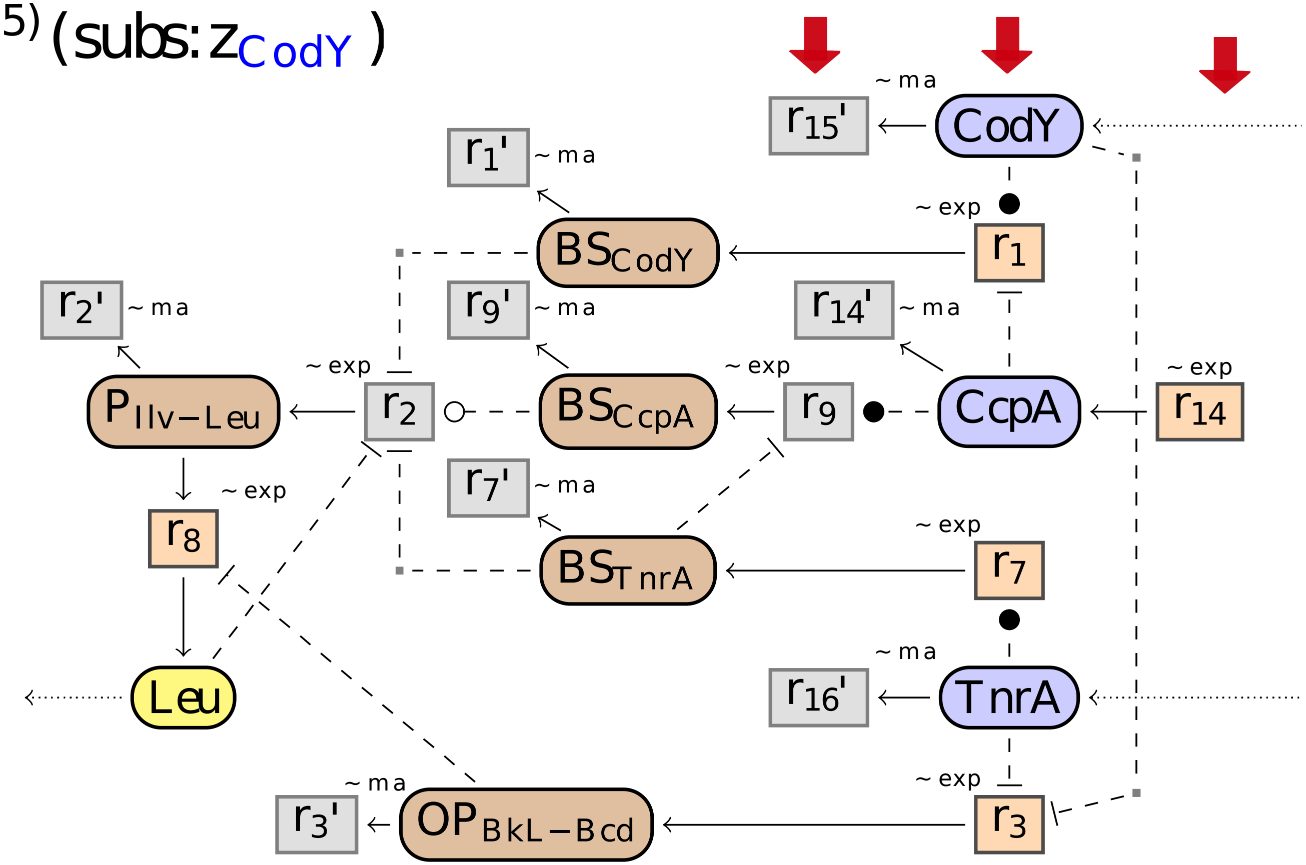
Qualitative reasoning

$x_{CodY} = v_{r_{15'}}$



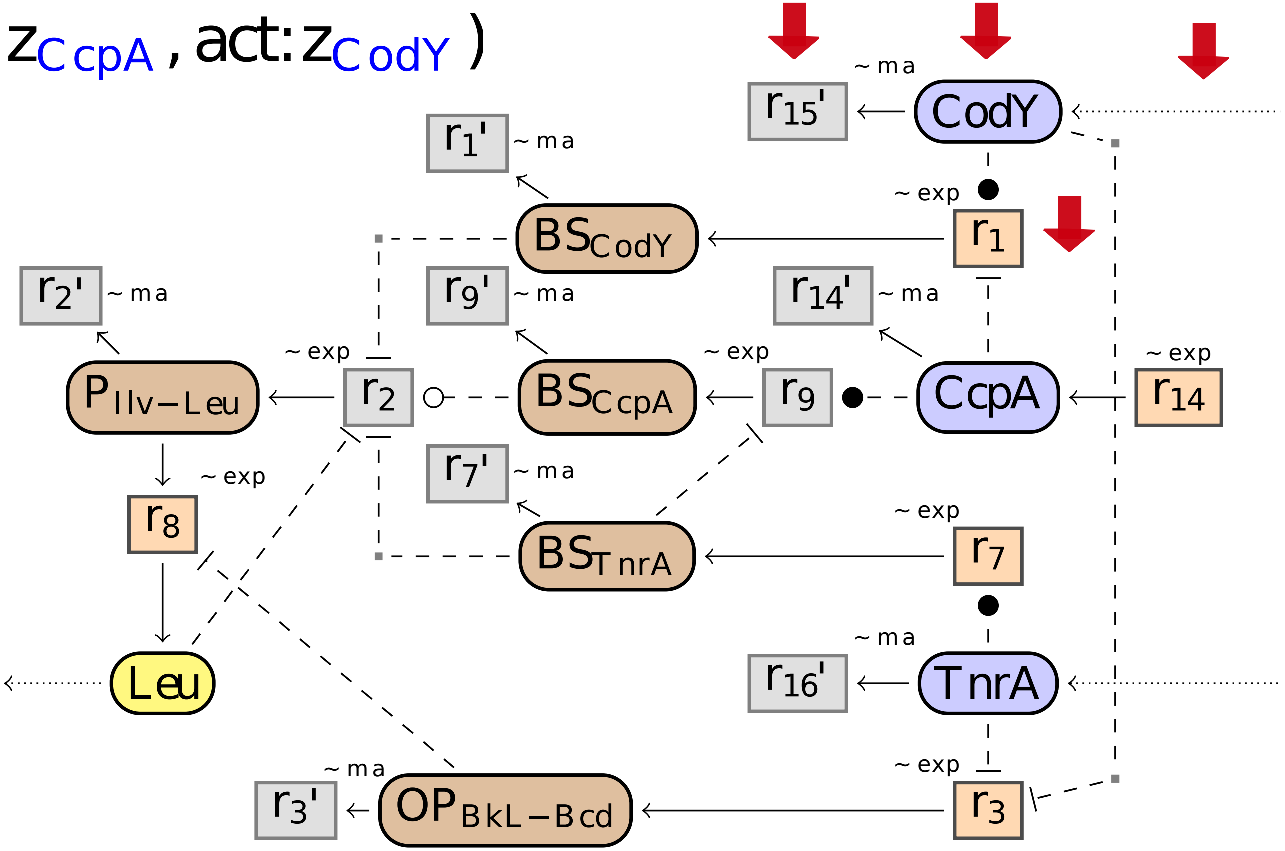
Qualitative reasoning

$$v_{r_{15}} = ma^{(15)}(subs: z_{CodY})$$



Qualitative reasoning

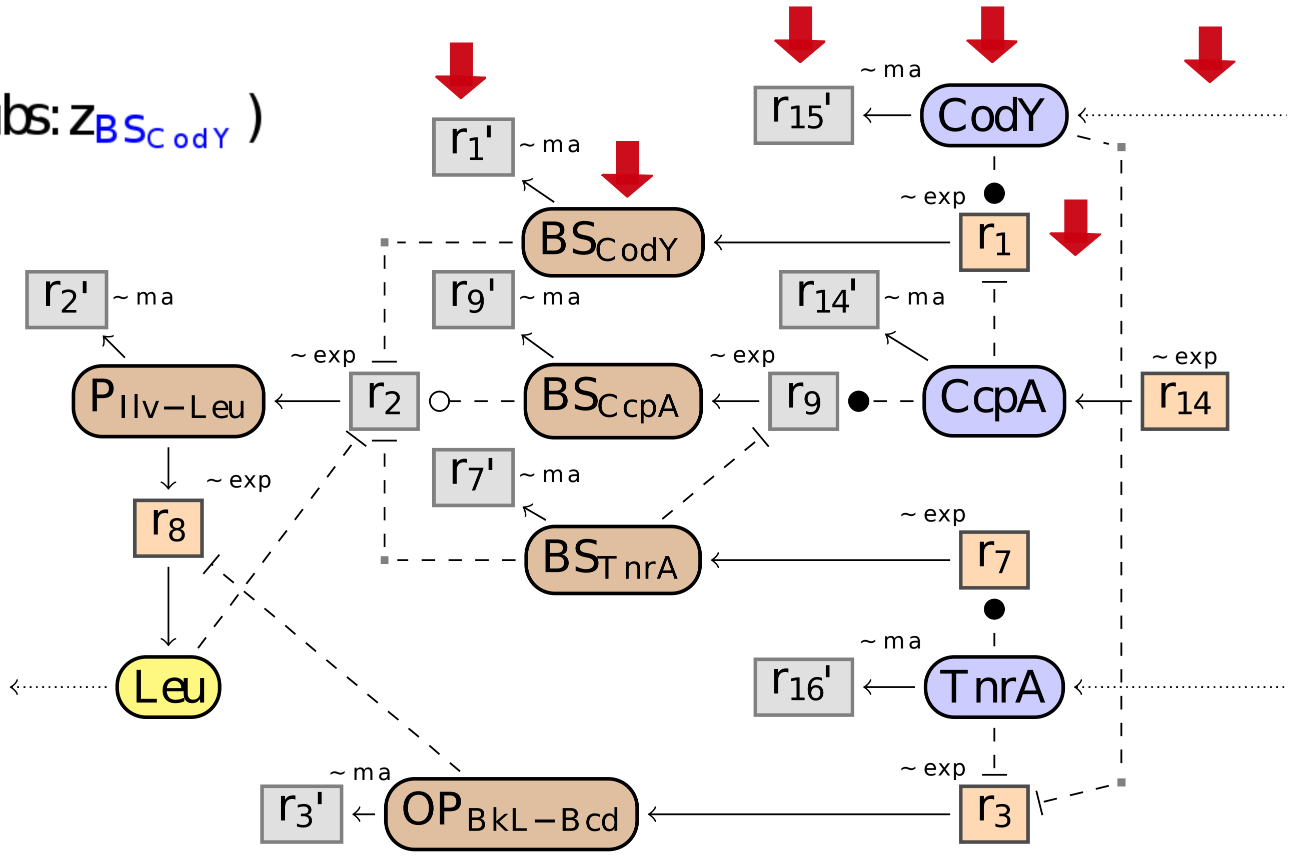
$$v_{r_1} = \exp^{(1)}(\mathrm{inh}: z_{\mathrm{CcpA}}, \mathrm{act}: z_{\mathrm{CodY}})$$



Qualitative reasoning

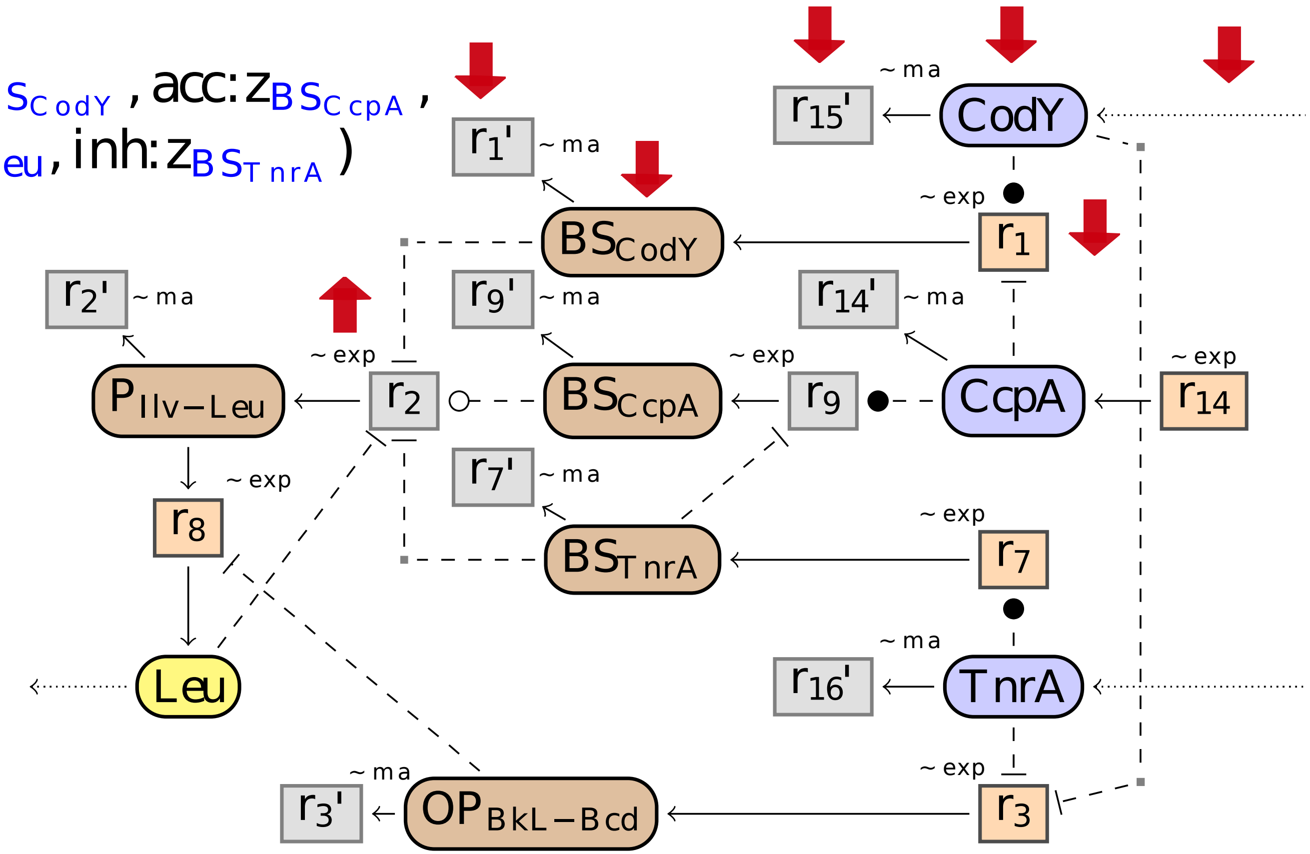
$$v_{r_1} = v_{r_{1'}}$$

$$v_{r_1} = ma^{(1)}(subs: z_{BS_{CodY}})$$



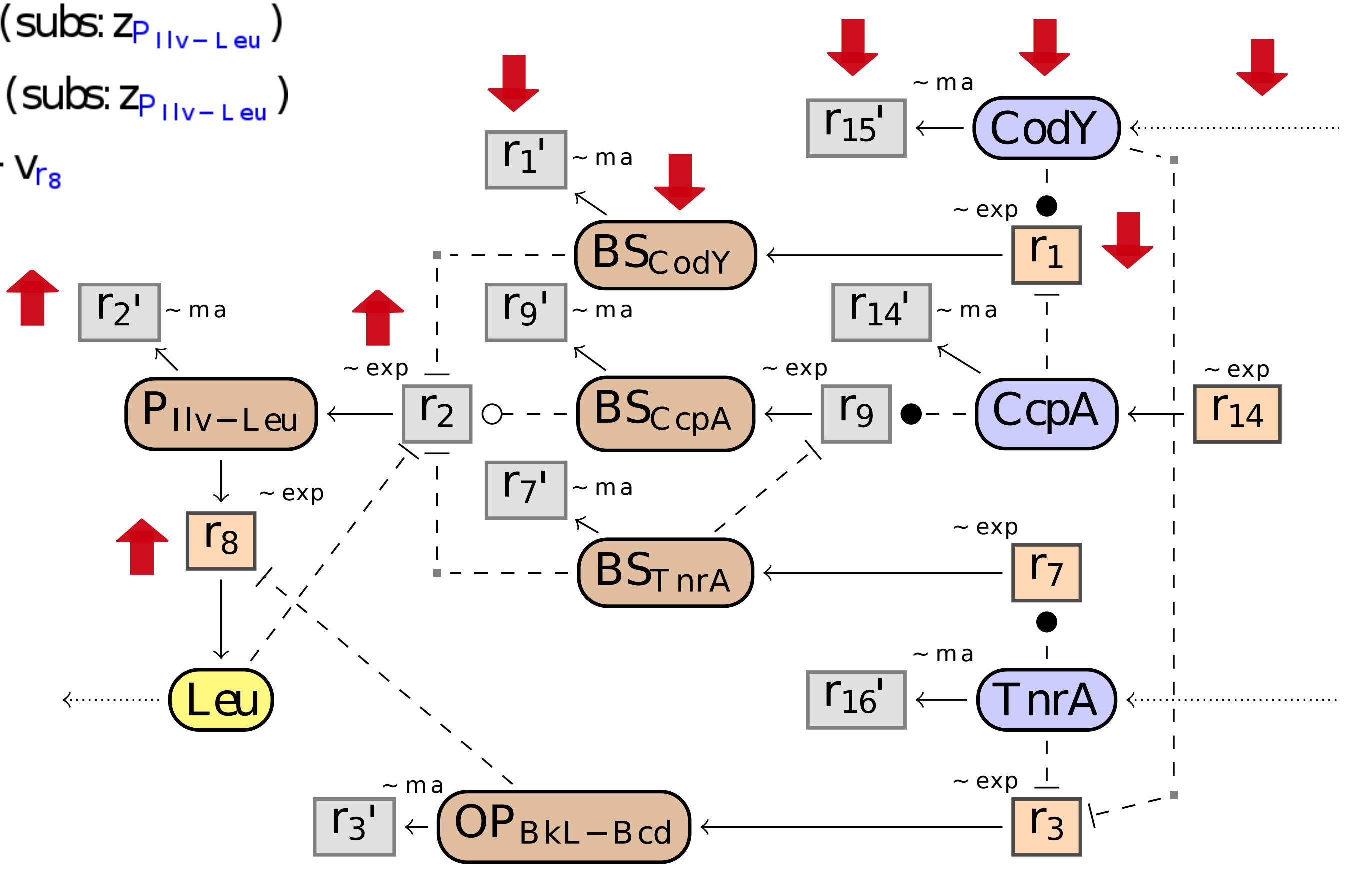
Qualitative reasoning

$$v_{r_2} = \exp^{(2)}(\text{inh}{:}z_{BS_{CodY}}, \text{acc}{:}z_{BS_{CcpA}}, \text{inh}{:}z_{Leu}, \text{inh}{:}z_{BS_{TnrA}})$$



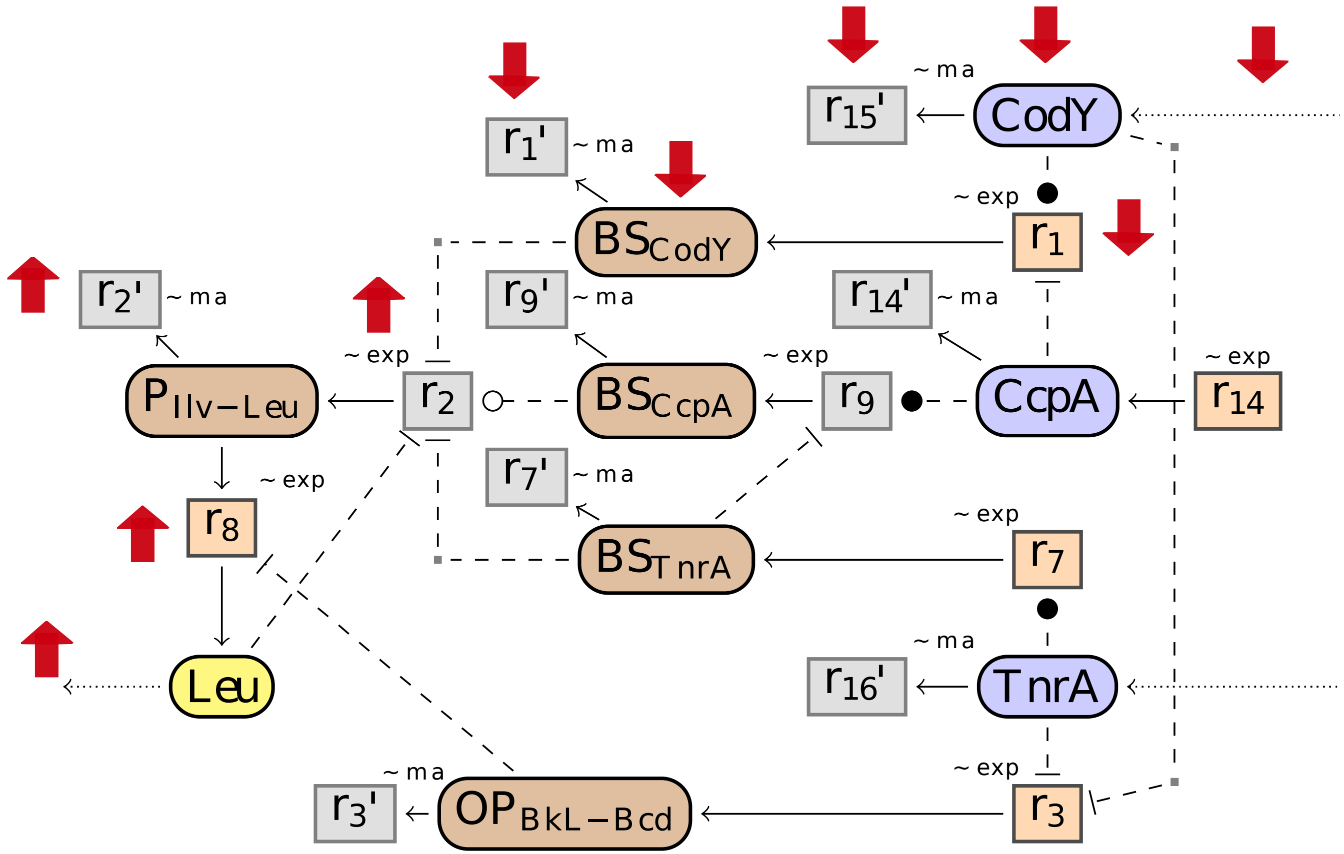
Qualitative reasoning

$$v_{r_{2'}} = ma^{(2)}(subs: z_{P_{Ilv-Leu}})$$
$$v_{r_8} = exp^{(8)}(subs: z_{P_{Ilv-Leu}})$$
$$v_{r_2} = v_{r_2}' + v_{r_8}$$



Qualitative reasoning

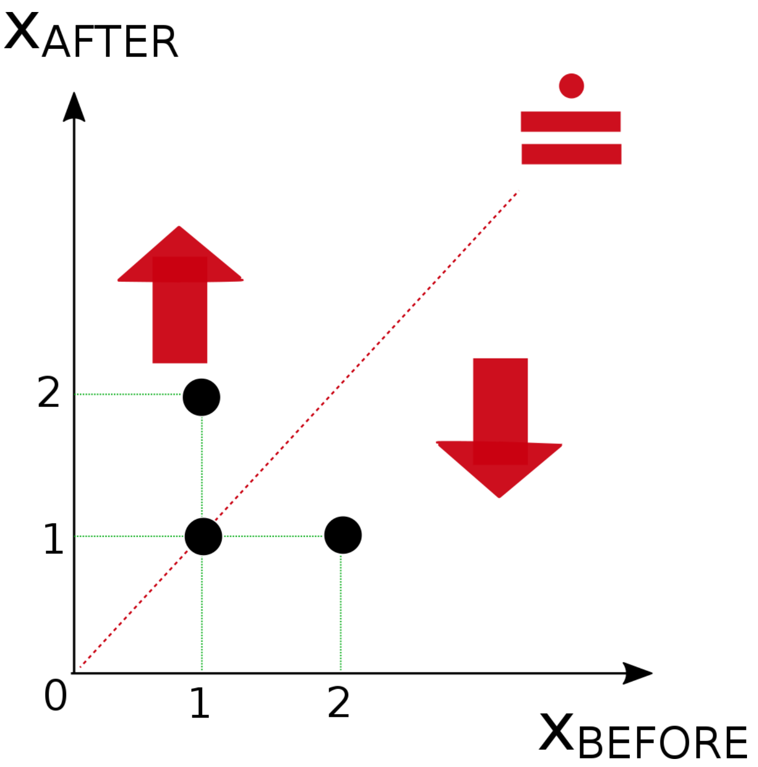
$$v_{r_8} = y_{Leu}$$



Abstract values

Variables:

- species concentrations
- reaction speeds



From $\mathbb{R}^+$ to $\Delta_3 = \{\uparrow, \doteq, \downarrow\}$

Abstract operators

- Operators redefined over finite domain Δ

abstract product

$$a \cdot^{\Delta} b$$

		b		
	$\cdot^{\Delta}$	$\uparrow$	$\doteq$	$\downarrow$
a	$\uparrow$	$\uparrow$	$\uparrow\doteq$	$\downarrow\doteq\uparrow$
	$\doteq$	$\uparrow\doteq$	$\doteq$	$\downarrow\doteq$
	$\downarrow$	$\downarrow\doteq\uparrow$	$\downarrow\doteq$	$\downarrow$

Abstract operators

- Operators redefined over finite domain Δ
  - from functions to relations

abstract product

$$a \cdot^{\Delta} b$$

			b	
	$\cdot^{\Delta}$	$\uparrow$	$\doteq$	$\downarrow$
a	$\uparrow$	$\uparrow$	$\uparrow\doteq$	$\downarrow\doteq\uparrow$
	$\doteq$	$\uparrow\doteq$	$\doteq$	$\downarrow\doteq$
	$\downarrow$	$\downarrow\doteq\uparrow$	$\downarrow\doteq$	$\downarrow$

Abstract constraints

$$(\mathrm{Leu}) \qquad v_{r_8} = y_{\mathrm{Leu}}$$

$$(P_{\mathrm{Ilv-Leu}}) \qquad v_{r_2} = v_{r_2} + v_{r_8}$$

$$v_{r_2} = \mathrm{exp}^{(2)}(\mathrm{inh}: z_{BS_{\mathrm{CodY}}}, \mathrm{acc}: z_{BS_{\mathrm{CcpA}}}, \mathrm{inh}: z_{\mathrm{Leu}}, \mathrm{inh}: z_{BS_{\mathrm{TnrA}}})$$

$$v_{r_8} = \mathrm{exp}^{(8)}(\mathrm{subs}: z_{P_{\mathrm{Ilv-Leu}}})$$

$$v_{r_2} = \mathrm{ma}^{(2)}(\mathrm{subs}: z_{P_{\mathrm{Ilv-Leu}}})$$

$$y_{\mathrm{Leu}} = \mathrm{ma}^{(0)}(\mathrm{subs}: z_{\mathrm{Leu}})$$

$$v_{r_2} \in v_{r_2} + y_{\mathrm{Leu}}$$

$$y_{\mathrm{Leu}} = z_{\mathrm{Leu}}$$

$$v_{r_2} \in z_{BS_{\mathrm{CcpA}}} \cdot \mathrm{Inh}(z_{BS_{\mathrm{CodY}}} + z_{\mathrm{Leu}} + z_{BS_{\mathrm{TnrA}}})$$

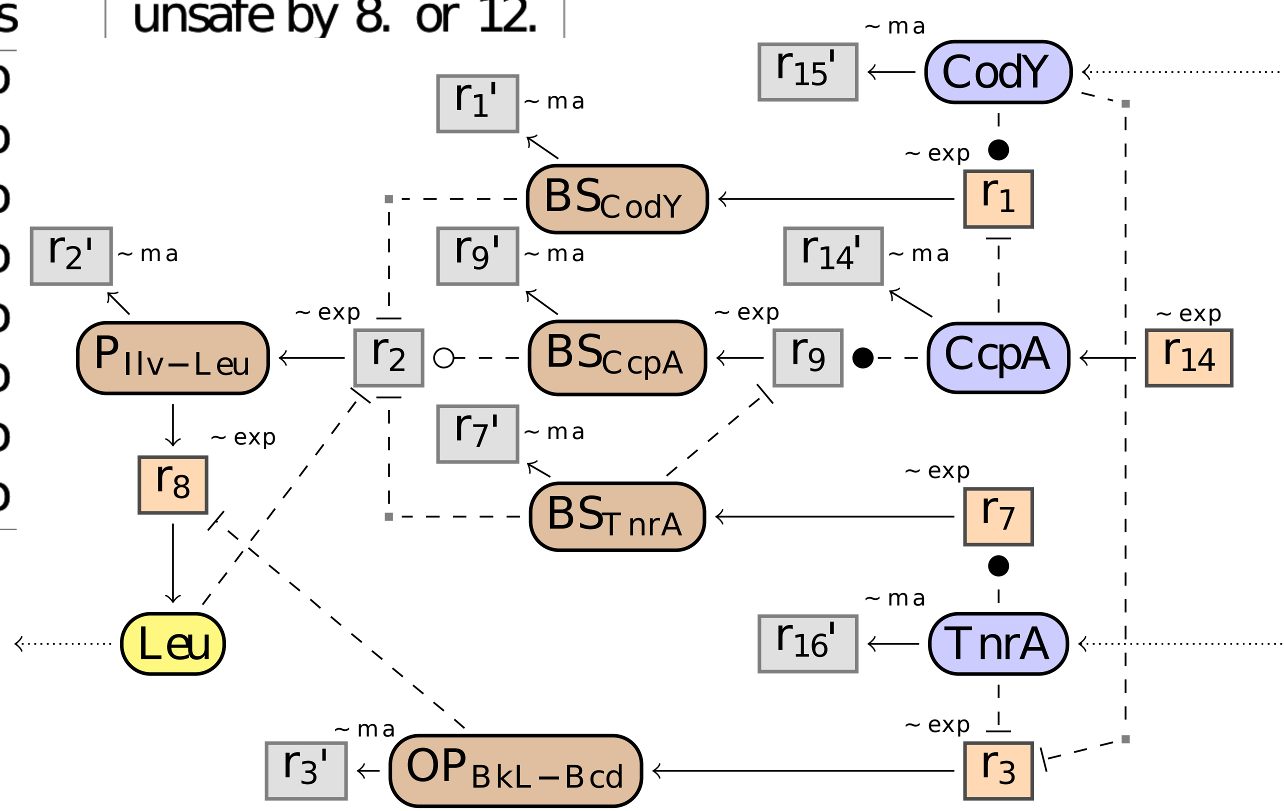
$$v_{r_2} = z_{P_{\mathrm{Ilv-Leu}}}$$

$$y_{\mathrm{Leu}} = z_{P_{\mathrm{Ilv-Leu}}}$$

Minizinc constraints solver

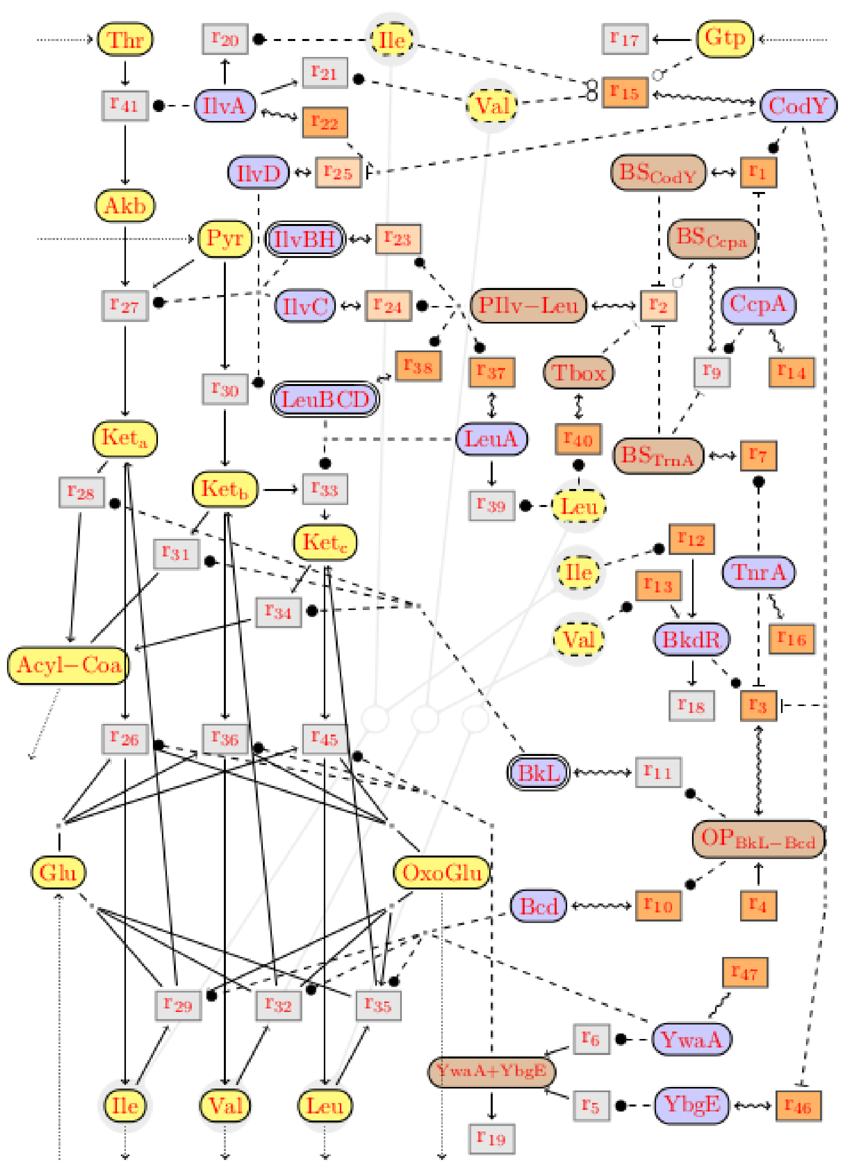
Prediction for PIlv promoter

	x_{CodY}	x_{TnrA}	y_{Leu}	target satisfied	quality
1	$\downarrow$	$\doteq$	$\uparrow$	yes	safe
2	$\doteq$	$\downarrow$	$\uparrow$	yes	safe
3	$\downarrow$	$\downarrow$	$\uparrow$	yes	safe
4	$\downarrow$	$\uparrow$	$\uparrow$	yes	unsafe by 7. or 11.
5	$\uparrow$	$\downarrow$	$\uparrow$	yes	unsafe by 8. or 12.
6	$\doteq$	$\doteq$	$\doteq$	no	
7	$\downarrow$	$\uparrow$	$\doteq$	no	
8	$\uparrow$	$\downarrow$	$\doteq$	no	
9	$\doteq$	$\uparrow$	$\downarrow$	no	
10	$\uparrow$	$\doteq$	$\downarrow$	no	
11	$\downarrow$	$\uparrow$	$\downarrow$	no	
12	$\uparrow$	$\downarrow$	$\downarrow$	no	
13	$\uparrow$	$\uparrow$	$\downarrow$	no	



(CMSB'15, BioSystem'16)

Predictions for Leucine Overproduction



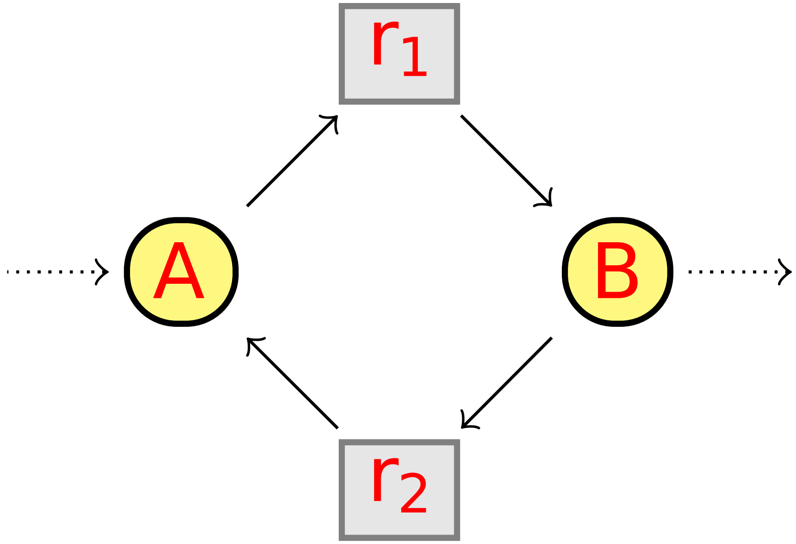
- 21 knockout candidates
- 12 single knockouts predicted
- 6 successfully tested experimentally
- knockout of CodY is best: 7 fold increase of surfactin

(BioTechnology Journal 2015)

Plan

- Background
  - Modeling language for reaction networks with partial kinetic information
  - Prediction by abstract interpretation
- **Over-approximation problem**
- Additional constraints from 'elementary modes'

Over-approximation problem

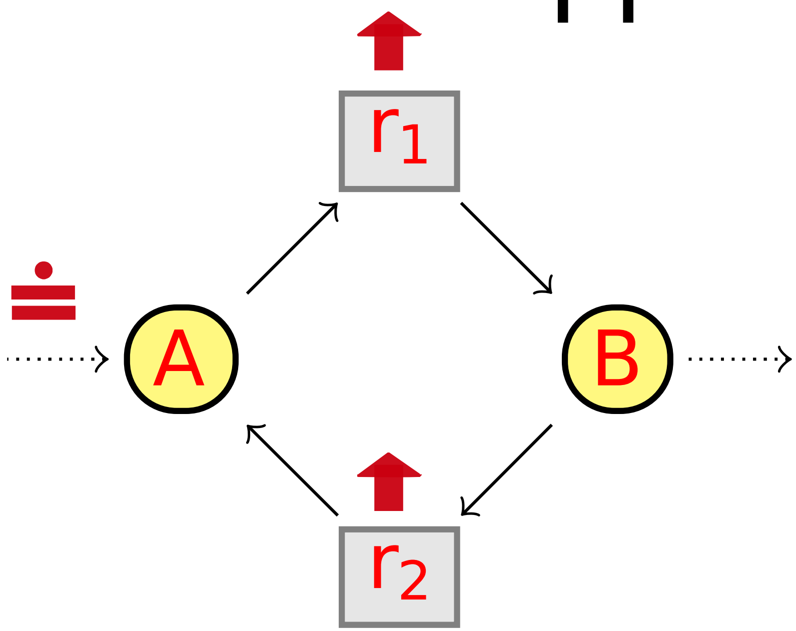


Information lost by abstracting

$$\begin{cases} +^{\Delta}(v_{r2}, x_A, v_{r1}) \\ +^{\Delta}(v_{r2}, y_B, v_{r1}) \end{cases}$$

$$\begin{cases} v_{r2} + x_A = v_{r1} \\ v_{r2} + y_B = v_{r1} \end{cases}$$

Over-approximation problem

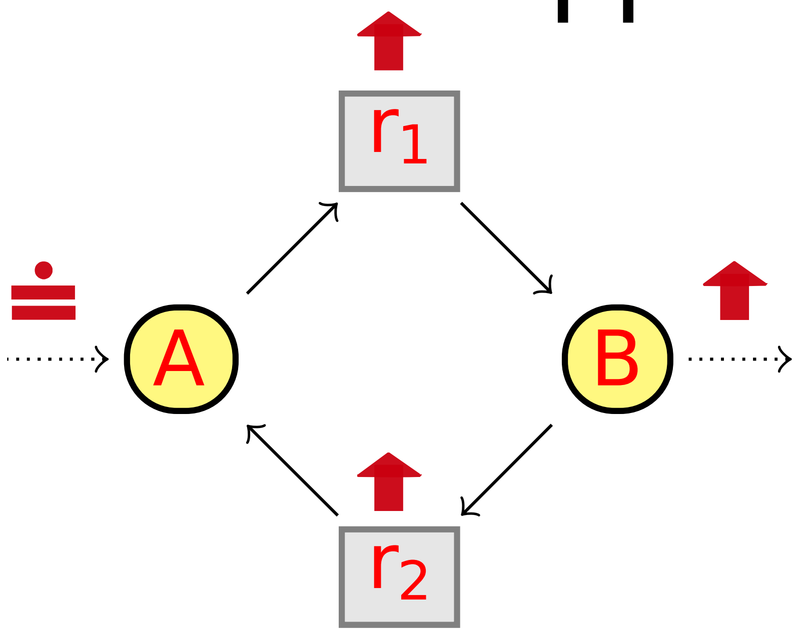


$+$	$\uparrow$	$\doteq$	$\downarrow$
$\uparrow$	$\uparrow$	$\uparrow$	$\downarrow\doteq\uparrow$
$\doteq$	$\uparrow$	$\doteq$	$\downarrow$
$\downarrow$	$\downarrow\doteq\uparrow$	$\downarrow$	$\downarrow$

$$\begin{cases} +^{\Delta}(v_{r2}, x_A, v_{r1}) \\ +^{\Delta}(v_{r2}, y_B, v_{r1}) \end{cases}$$

$$\begin{cases} v_{r2} + x_A = v_{r1} \\ v_{r2} + y_B = v_{r1} \end{cases}$$

Over-approximation problem



$+$	$\uparrow$	$\doteq$	$\downarrow$
$\uparrow$	$\uparrow$	$\uparrow$	$\downarrow\doteq\uparrow$
$\doteq$	$\uparrow$	$\doteq$	$\downarrow$
$\downarrow$	$\downarrow\doteq\uparrow$	$\downarrow$	$\downarrow$

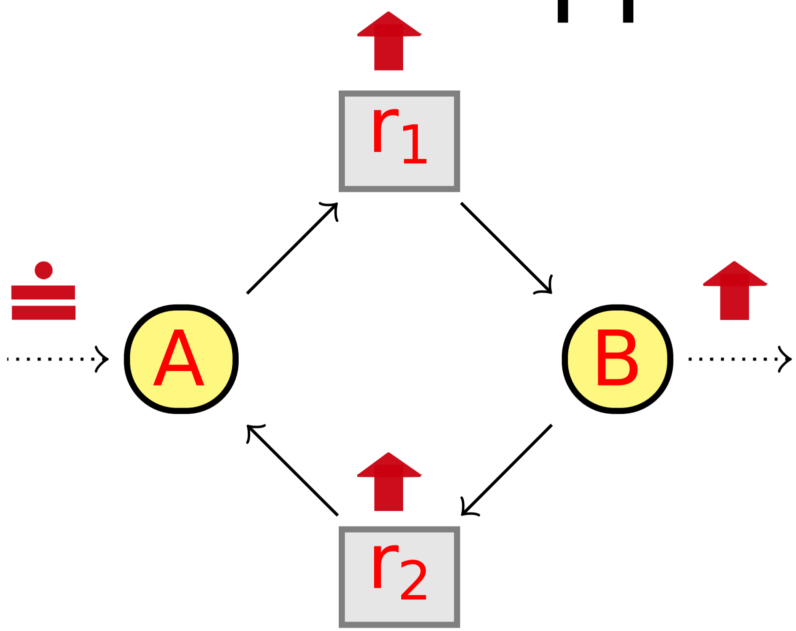
$$\begin{cases} +^{\Delta}(v_{r2}, x_A, v_{r1}) \\ +^{\Delta}(v_{r2}, y_B, v_{r1}) \end{cases}$$

MISSING

$$\left\{ =^{\Delta} (x_A, y_B) \right.$$

$$\begin{cases} v_{r2} + x_A = v_{r1} \\ v_{r2} + y_B = v_{r1} \end{cases}$$

Over-approximation problem



$+$	$\uparrow$	$\doteq$	$\downarrow$
$\uparrow$	$\uparrow$	$\uparrow$	$\downarrow\doteq\uparrow$
$\doteq$	$\uparrow$	$\doteq$	$\downarrow$
$\downarrow$	$\downarrow\doteq\uparrow$	$\downarrow$	$\downarrow$

$$\begin{cases} +^{\Delta}(v_{r2}, x_A, v_{r1}) \\ +^{\Delta}(v_{r2}, y_B, v_{r1}) \end{cases}$$

$$\{ =^{\Delta} (x_A, y_B)$$

Additional constraint

$$\begin{cases} v_{r2} + x_A = v_{r1} \\ v_{r2} + y_B = v_{r1} \end{cases}$$

$$\boxed{x_A = y_B}$$

IDEA

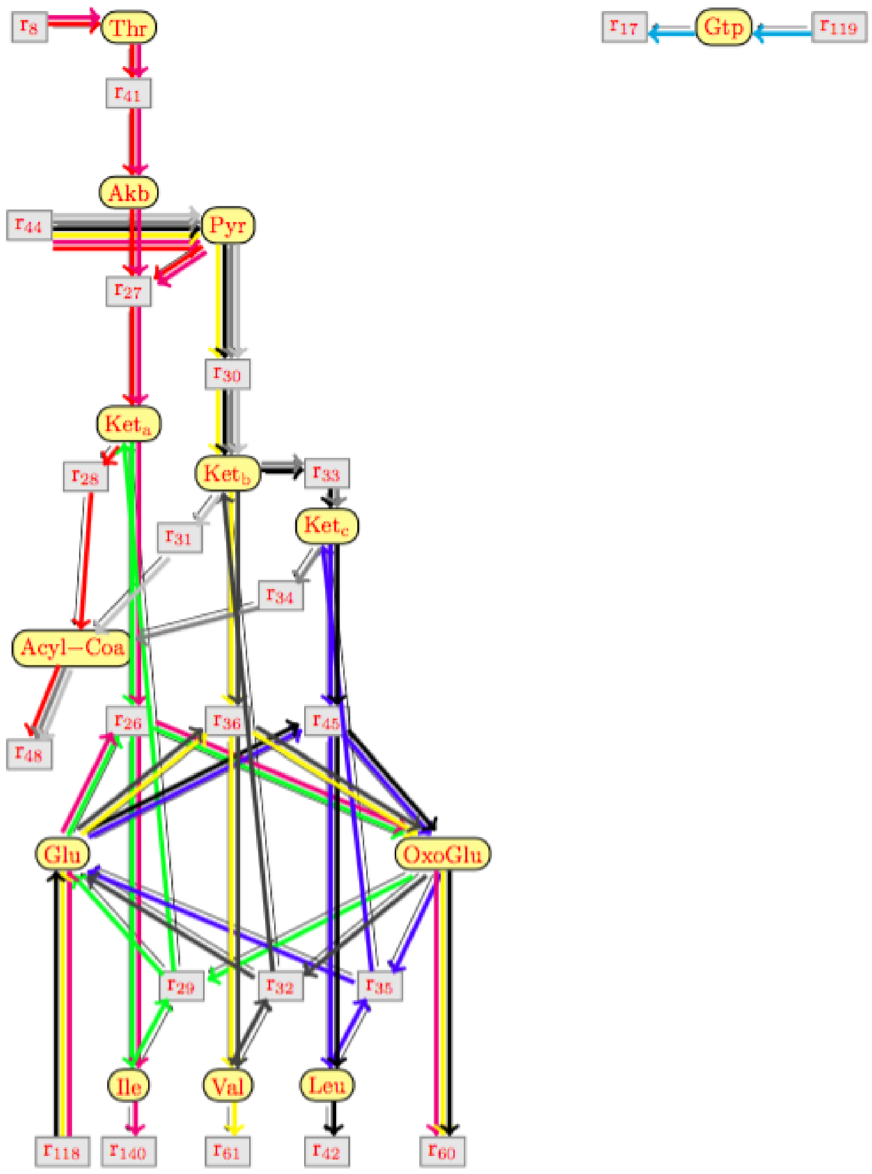
Plan

- Background
  - Modeling language for reaction networks with partial kinetic information
  - Prediction by abstract interpretation
- Over-approximation problem
- **Additional constraints from 'elementary mode'**

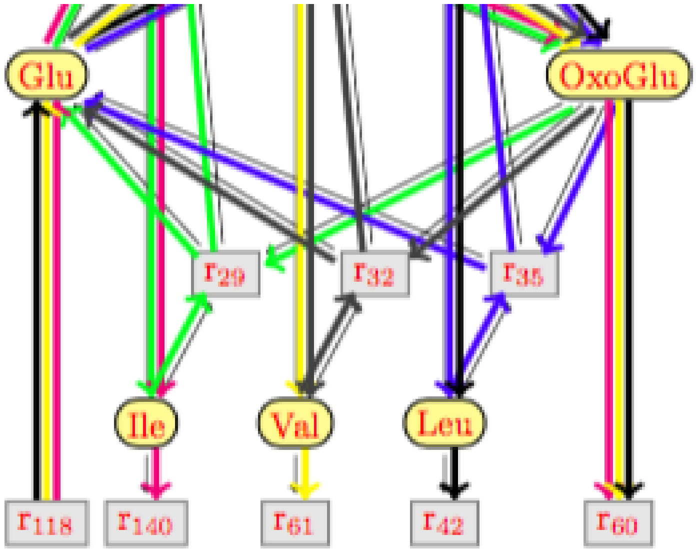
Partial solution

Elementary Flux Mode (EFM)

Minimal 'subset' of reactions that is steady



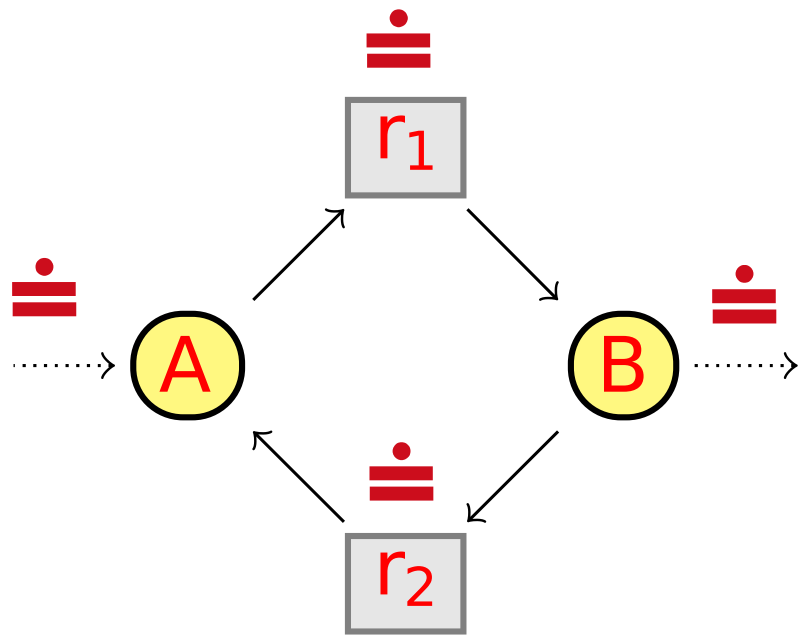
Partial solution



$$v_{r118} = v_{r60}$$

Equality of speed of reactions by their composition in terms of EFM

Result



- 👍 fewer solutions = predictions
- 👍 more safe solutions
- 👎 higher computation time

Current work

- Partial solution not sufficient
- All consequences of the arithmetic system

  → Infinite number !
- Finite set of abstract constraint exists !

Is it possible to deduce this set of abstract constraints automatically ? And How ?

Future work

- Predictions of multiple knockouts
  - Heuristics for improving constraint solver
- Quantitative reasoning
- Medical applications?

Thank you!

But we want ...

$$\begin{cases} v_{r2} + x_A = v_{r1} \\ v_{r2} + y_B = v_{r1} \end{cases} \qquad x_A = y_B$$

All linear combinations of the rows

of the stoichiometry matrix = **row space**

→ infinite number of equations !

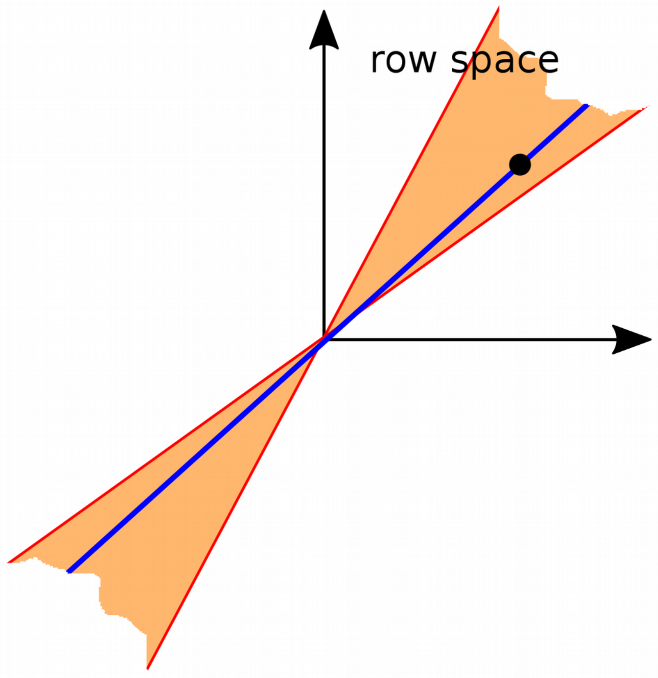
Finite set of abstract constraints ?

… one point along each direction...

$$x_A = x_B$$

$$2x_A = 2x_B$$

$$\Rightarrow \; =^{\Delta} (x_A, x_B)$$



We do not need equivalent equations

Finite set yes !

m^n linear combination of variables

with :

- n variables
- m card(domain)

But which one are sufficient ?